

Nicola Scafetta¹, Tim Imholt², Paolo Grigolini^{1,3,4}, and Jim Roberts².¹Center for Nonlinear Science, University of North Texas, P.O. Box 311427, Denton, Texas 76203-1427²Center for Climate Analysis, University of North Texas, P.O. Box 311427, Denton, Texas 76203-1427³Dipartimento di Fisica dell'Università di Pisa and INFN, Piazza Torricelli 2, 56127 Pisa, Italy⁴Istituto di Biofisica CNR, Area della Ricerca di Pisa, Via Alfieri 1, San Cataldo 56010 Ghezzano-Pisa, Italy

(February 6, 2002)

This paper presents a wavelet multiresolution analysis of a time series dataset to study the correlation between the real temperature data and three temperature model reconstructions at different scales. We show that the Mann *et al.* model reconstructs the temperature better at all temporal resolutions. We show and discuss the wavelet multiresolution analysis of the Mann's temperature reconstruction for the period from 1400 to 2000 A.D.E.

I. INTRODUCTION

The study of past climate trends, or variability is absolutely imperative if any kind of prediction is to be made about the future of our global environment, [1]. With the goal of understanding how our climate has changed in the past, several different models have been developed from various Paleoclimatology proxies, [2]. These proxy records, located all over the globe, are inherent in reconstruction temperature trends from periods before instrumentation records are available. These longer temperature records are necessary due to the short span of actual temperature records taken via instrumentation. The longer the time series of temperature data available the better a long term prediction of future temperature behavior can be made. These proxy data sets include information recovered from the records of tree rings, pollen samples, ice cores, corals, marine sediments as well as others. The models constructed, in order to be more accurate, take into account many different proxies in determining past climate trends. This paper makes a comparison of these climate models with instrumentation readings in a new way. The results show that the model produced by Mann *et al.* [3] gives a slightly better fit to known temperature values than several of the other paleoclimate models.

II. TEMPERATURE MODELS

There are three climate models compared in this paper, Fig. 1. The comparisons are made to known temperature anomaly data taken via instrumentation and reported by Jones *et al.* [4]. This data set is largely considered as the most accurate global temperature data set known. The models compared are those of Mann *et al.* [3], [5] Briffa *et al.* [6], [7] and Jones *et al.* [8], [9] These models all contain at a minimum data from the years 1400 A.D.E. to 1980 A.D.E. (All dates from this point are to be considered

A.D.E.)

The Jones *et al.* [8], [9] dataset was generated by averaging 17 temperature reconstructions from both hemispheres. These sites, once averaged together, form a dataset 1000 years long with a temporal resolution of 1 year. The proxy data types utilized were tree rings, ice cores, corals and historical documents. The Mann *et al.* [3], [5] data was a reconstruction from sites globally and extends from the years 1400-1980. The reconstruction was performed utilizing proxy data from tree rings, ice core, ice melt, long historical records (from Bradley and Jones [10]), coral, and long instrumentation records. The Briffa *et al.* [6] [7] dataset extends from 1400-1994 and was a reconstruction utilizing only tree ring data.

III. WAVELET MULTIREOLUTION ANALYSIS

Wavelet analysis [11] is a new powerful method to analyze time series. Wavelet Transform makes use of scaling functions, the wavelets, that have the characteristics of being localized in space and in frequencies. These functions must integrate to zero and the square of them must integrate to unity. A scaling coefficient τ characterizes a wavelet. The length 2τ measures the width of the wavelet. Two typical wavelets widely used in the continuous wavelet transform are the Haar wavelet and the Mexican hat wavelet [11]. The Haar wavelet is defined by

$${}^{(H)}\tilde{\psi}_{\tau,t}(u) \equiv \begin{cases} -1/\sqrt{2\tau}, & t - \tau < u < t \\ 1/\sqrt{2\tau}, & t < u < t + \tau \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The Mexican hat wavelet is the second derivative of a Gaussian. The length 2τ defines the scale analyzed by the wavelet. Given a signal $\xi(u)$, the Continuous Wavelet Transform is defined by

$$W(\tau, t) = \int_{-\infty}^{\infty} \tilde{\psi}_{\tau,t}(u) \xi(u) du . \quad (2)$$

The original signal can be recovered from its Continuous Wavelet Transform via

$$\xi(u) = \frac{1}{C_{\tilde{\psi}}} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W(\tau, t) \tilde{\psi}_{\tau,t}(u) dt \right] \frac{d\tau}{\tau^2} . \quad (3)$$

The double integral of Eq. (3) suggests that the original signal may be decomposed in “continuous details” that depend on the scale coefficient τ . However, it is not easy to handle the results of the Continuous Wavelet Transform because of the continuous nature of the decomposition. There exists a discrete version of the wavelet transform, the Maximum Overlap Discrete Wavelet Transform (MODWT), which is the basic tool needed for studying time series of N data via wavelet. In the book of Percival and Walden, “*Wavelet Methods for Time Series Analysis* [11], the reader can find all mathematical details. For the purpose of this paper, it is important to have in mind only one of the important properties of the MODWT, the Wavelet Multiresolution Analysis (WMA). It is possible to prove that given an integer j_0 such that $2^{j_0} < N$, where N is the number of the data, the original time series represented by the vector \mathbf{X} can be decomposed in:

$$\mathbf{X} = S_{J_0} + \sum_{j=1}^{J_0} D_j , \quad (4)$$

with

$$S_{j-1} = S_j + D_j . \quad (5)$$

The detail D_j represents changes on a scale of $2\tau = 2^j$, while the smooth S_{J_0} represents averages of a scale of $2\tau = 2^{J_0}$. To better appreciate the WMA, Fig. 2 shows the WMA of the temperature in the years 1860-2000. The analysis is done by using the Daubechies *least asymmetric* scaling wavelet filter (LA8) [11]. LA8 wavelet look similar to the Mexican hat but they are asymmetric. We plot the WMA for $J_0 = 4$. Fig. 2a shows the temperature (dashed line) and the smooth average S_4 (solid line). Figs. 2b-e show the details D_4, D_3, D_2, D_1 . Detail D_4 shows the presence of irregular oscillations with a period of almost 20-22 years. Detail D_3 shows the presence of more regular oscillations with a period of almost 10-11 years. Details D_2 and D_1 show the fluctuations of the temperature at scale $2\tau = 2^2 = 4$ and $2\tau = 2^1 = 2$ years respectively. According to Eqs. (4) and (5), the sum of all four details and the smooth average S_4 give the original signal. The smooth average S_3 is given by $S_4 + D_4$. The smooth averages S_2 and S_1 can be obtained by summing first D_3 and then D_2 to S_3 respectively.

IV. MULTIREOLUTION CORRELATION ANALYSIS

The use of the Multiresolution Correlation Analysis via wavelet is a simple procedure. We decompose the temperature data and the three temperature reconstructions by Briffa, Jones and Mann by using WMA introduced in the previous section. Then we evaluate the linear correlation coefficient r between the real temperature data and each of the reconstructions for each smooth average and each detail. For pairs of quantities $(x_i; y_i)$; $i = 1, \dots, N$, the linear correlation coefficient r is given by the formula

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} , \quad (6)$$

where, as usual, \bar{x} is the mean of the x , \bar{y} is the mean of the y . The value of r lies between -1 and 1, inclusive. It takes on a value of 1, termed “completely positive correlation,” when the data points lie on a perfect straight line with positive slope, with x and y increasing together. The value 1 holds independent of the magnitude of the slope. If the data points lie on a perfect straight line with negative slope, y decreasing as x increases, then r has the value -1; this is called “completely negative correlation.” A value of r near zero indicates that the variables x and y are uncorrelated. When a correlation is known to be significant, r is one conventional way of summarizing its strength. Because we want to determine which temperature model reproduces better the real temperature at each scale, the best model is the one that gives a linear correlation coefficient r closest to 1. Table 1 shows the correlation coefficient between the real temperature data and each of the reconstruction for each smooth average and each detail.

Table 1 shows clearly that Mann’s model is the best one in all cases, because its linear correlation coefficient is the closest to 1 for each analysis. Briffa’s model is the worst one. However, if we analyze the details, Briffa’s model is better than Jones’ model. This is information that a simple look at Fig. 1 could not determine. Moreover, the linear correlation coefficients for the smooth averages are closer to 1 than the correlation coefficients concerning the details. This suggests that the models are made in such a way to reproduce better the smooth averages than the details. Finally, we observe that the worst correlation is for the details D_3 in all cases. This is very interesting because, as shown in Fig. 2c, the details D_3 represents changes on a scale of $2\tau = 2^3 = 8$ years that evidence the 10-11 years temperature periodicity that is connected with the 11 years solar cycle. The low value of r for the details D_3 suggests that the models give results uncorrelated to real data because of a random shifting. Fig 3 shows the comparison between the details D_4, D_3, D_2 and D_1 from Wavelet Multiresolution Analysis of the temperature reconstruction by Mann’s model (solid line)

and the real temperature (dashed line) during the year period (1856 - 1980). Figs. 3a, 3c and 3d show that the details D_4 , D_2 and D_1 of the temperature reconstruction afforded by Mann's model are satisfactorily correlated to the corresponding properties of the data concerning real temperature. As for D_3 , illustrated in Fig. 3b, we notice a kind of random shift between reconstruction and real temperature curve. In some time regions the maxima and minima of the reconstruction curve occur earlier, and in other time regions they occur later. However, the period and the amplitude of the 10-11 year cycle seem to be the same in both cases.

V. MANN'S MODEL ANALYSIS

In the previous section we proved that the Mann's model is the one that produces the most accurate temperature reconstruction of the real temperature data. The quality of agreement between temperature reconstruction and real temperature becomes unsatisfactory only for the wavelet detail D_3 , a detail of some importance since it refers to the time scale $2\tau = 2^3 = 8$ years, which is very close to the 10-11 year solar cycle. Figs. 4 show the WMA of the Mann's temperature reconstruction in the period 1400-1980 for $J_0 = 4$. Fig. 4a shows the Mann temperature model (dashed line) and the smooth average S_4 (solid line). The detail D_4 of Fig. 4b shows that the fastest changes on a scale of $2\tau = 2^4 = 16$ years happened during the periods 1485-1530, 1615-1660, 1685-1710, 1760-1790 and 1810-1840. The detail D_1 of Fig. 4c shows that the changes on a scale of $2\tau = 2^1 = 2$ year during the period 1700-1980 look stronger than those during the period 1400-1700. Further research is needed to determine whether the effect observed in the detail D_1 is natural or is due to a lack of precision of the data used to reconstruct the temperature. Finally, Fig. 5 shows the smooth average S_3 (solid line) for Mann's temperature reconstruction in the period 1400-1980. The dashed line is the real temperature during the years 1856-2000. The smooth average S_3 is the best smooth reconstruction of the temperature that can be done. In fact in Section 4, we proved that the correlation between the real temperature and the temperature reconstruction fails for details D_3 . Therefore, it is not realistic to include those details, as well as the details D_2 and D_1 , in a plausible smooth reconstruction of the temperature. However, in accordance again with the results of Section 4, the details D_2 and D_1 may be considered good enough for studying the changes on the scales of $2\tau = 2^2 = 4$ and $2\tau = 2^1 = 2$ years, respectively. Fig. 5 shows that the coldest period occurred during the years 1455-1475. From 1500 to 1920 the temperature had little fluctuation, only slightly more than 0.1 degree. There was a warm period from 1765-1778. Since 1920 the temperature had the highest growth in the last six centuries, almost 0.7 degree. The

temperature decreased slightly during the period 1945-1975 and then it started to increase again during the last 25 years.

VI. CONCLUSION

The conclusions clearly show that, at a minimum, wavelet analysis is a useful technique for comparison of these datasets. The superior dataset appears to be the Mann *et al.* [3], [5]. Future work will include utilizing this wavelet analysis technique in attempting to improve the existing models or in constructing a new paleoclimate model so that a better understanding of current and future climate behavior may be obtained.

-
- [1] P.D. Jones, T.J. Osborn, K.R. Briffa, The Evolution of Climate Over the Last Millennium, *Science* 292, 2001, 662-667.
 - [2] Stokstad, Eric, Myriad Ways to Reconstruct Past Climate, *Science* 292, 2001, 658-659.
 - [3] Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes, 1998, Global Scale Temperature Patterns and Climate Forcing over the past Six Centuries, *Nature* 392, 779-787 (1998).
 - [4] Jones, P.D., D.E. Parker, T.J. Osborn, and K.R. Briffa. 2001. Global and hemispheric temperature anomalies - land and marine instrumental records. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.
 - [5] Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes, 1998, Global Six Century Temperature Patterns, IGBP PAGES/World Data Center-A for Paleoclimatology Data Contribution Series 1998-016. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA.
 - [6] Briffa, K.R., P.D. Jones, F.H. Schweingruber and T.J. Osborn (1998) Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature* 393, 450-455.
 - [7] Briffa, *et al.*, 1998, Northern Hemisphere Temperature Reconstructions. IGBP PAGES/World Data Center-A for Paleoclimatology Data Contribution Series 98-022. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA.
 - [8] P.D. Jones, K.R. Briffa, T.P. Barnett, and S.F.B. Tett, 1998, High-resolution Paleoclimatic Records for the last Millennium: Interpretation, Integration and Comparison with General Circulation Model Control-run Temperatures, *The Holocene* 8, 1998, 455-471.
 - [9] P.D. Jones, K.R. Briffa, T.P. Barnett, and S.F.B. Tett, 1998, Millennial Temperature Reconstructions. IGBP PAGES/World Data Center-A for Paleoclimatology Data

Contribution Series 1998-039. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA.

- [10] Bradley, R.S. and Jones, P.D. 'Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends. *The Holocene* 3, 367-376 (1993).
- [11] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge (2000).

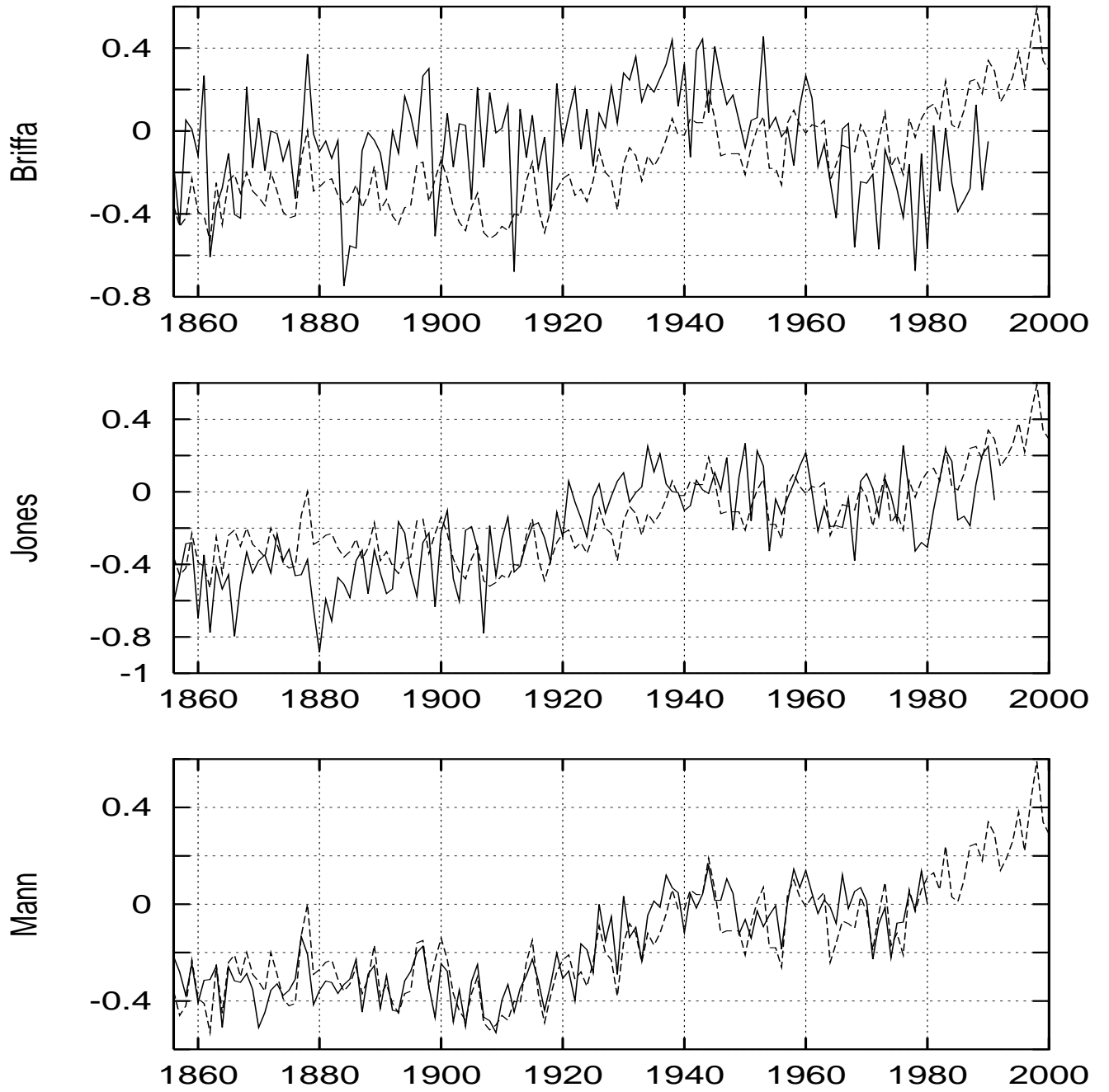


FIG. 1. Temperature reconstruction (1860:2000): Briffa, Jones and Mann models. The dashed line is the real temperature.

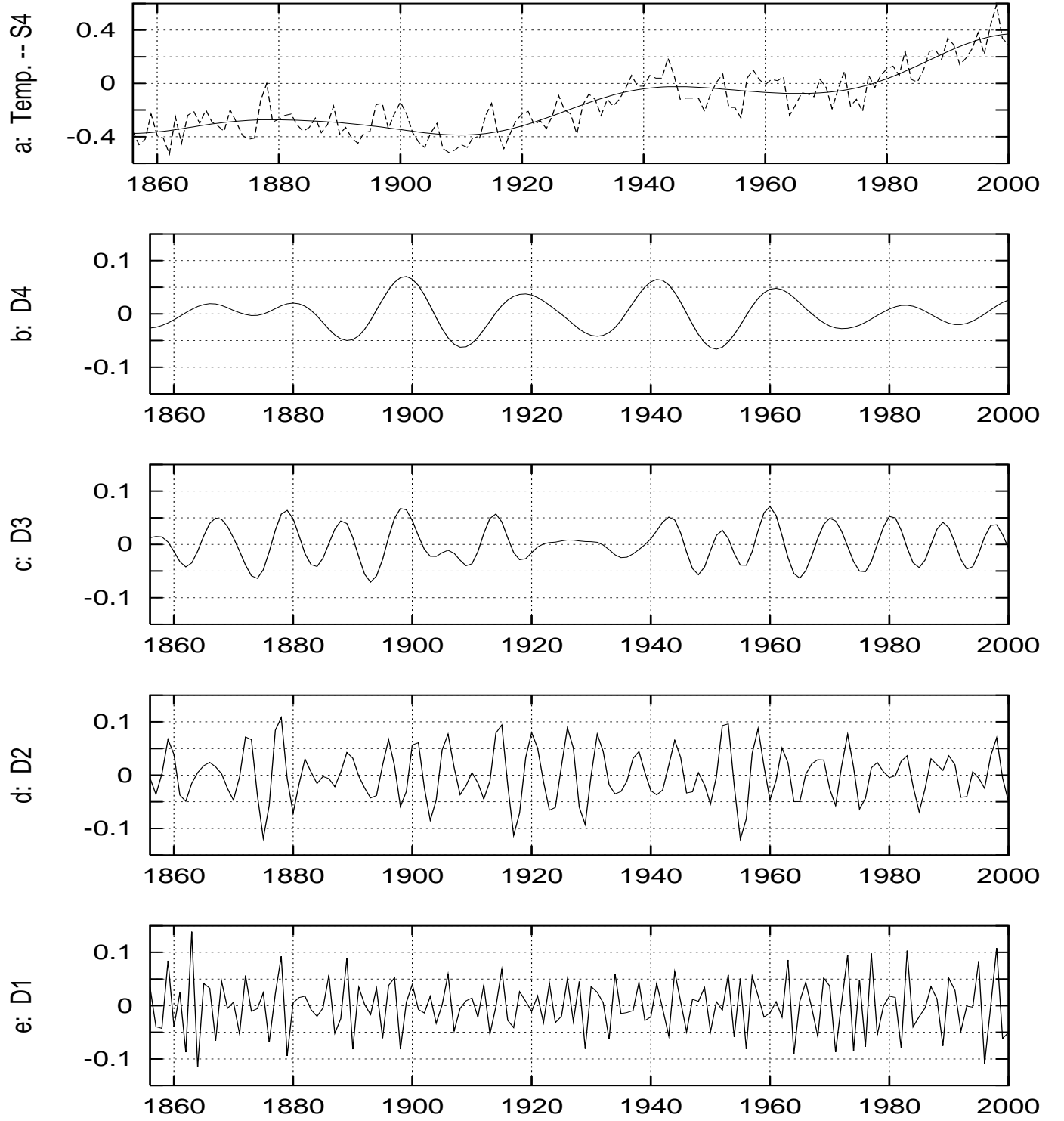


FIG. 2. Wavelet Multiresolution Analysis $J_0 = 4$ of the temperature. Fig. 2a shows the temperature (dashed line) and the smooth average S_4 (solid line). Figs. 2b-e show the details D_4 , D_3 , D_2 , D_1 .

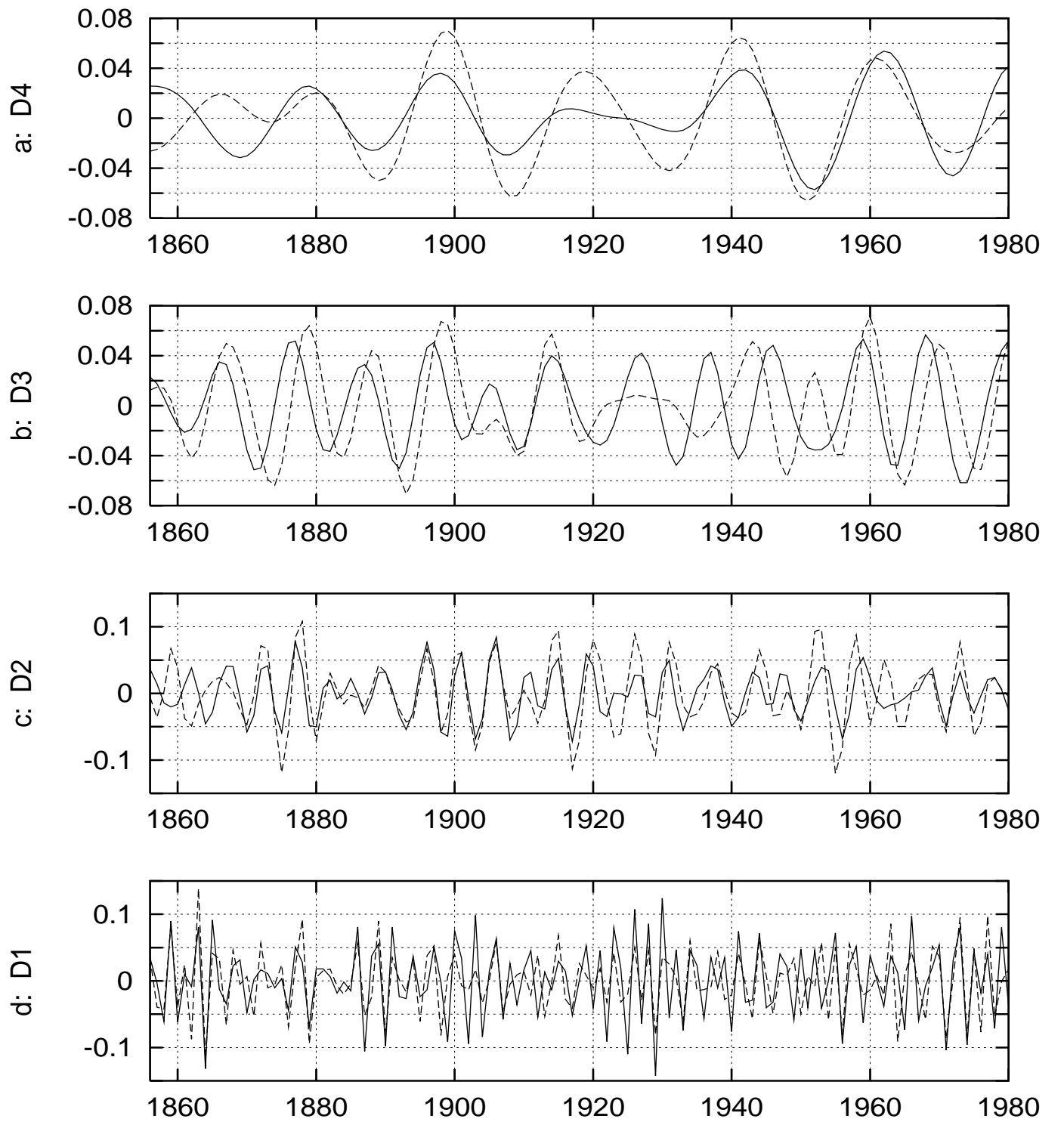


FIG. 3. Comparison between Temperature reconstruction according to Mann's model (solid line) and real temperature (dashed line). We show the details $D4$, $D3$, $D2$ and $D1$ obtained in both cases with the method of Wavelet Multiresolution Analysis: year (1856:1980).

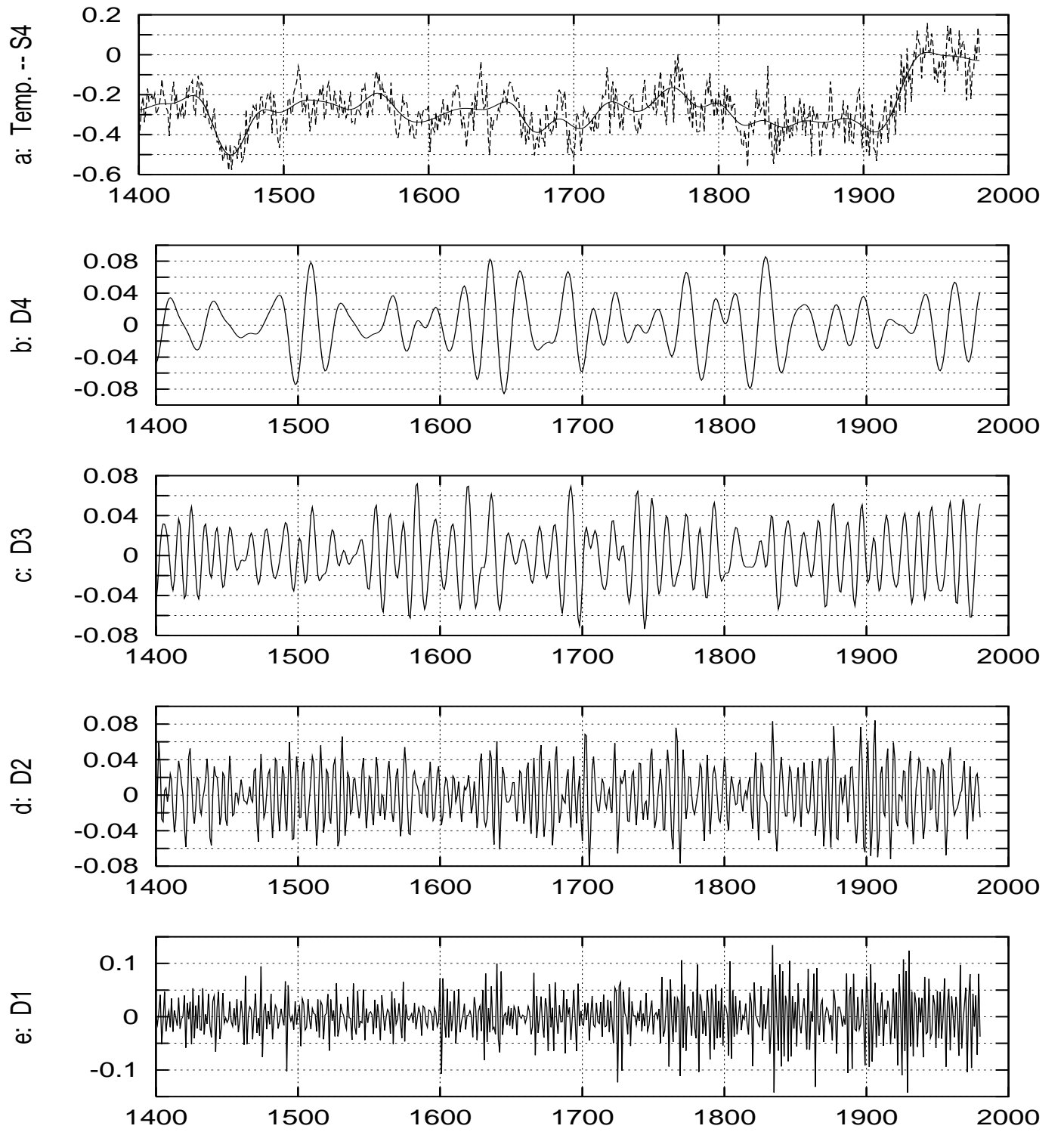


FIG. 4. Wavelet Multiresolution Analysis $J_0 = 4$ of the Mann temperature model; years 1400-1980. Fig. 4a shows the temperature model (dashed line) and the smooth average S_4 (solid line). Figs. 4b-e show the details D_4 , D_3 , D_2 , D_1 .

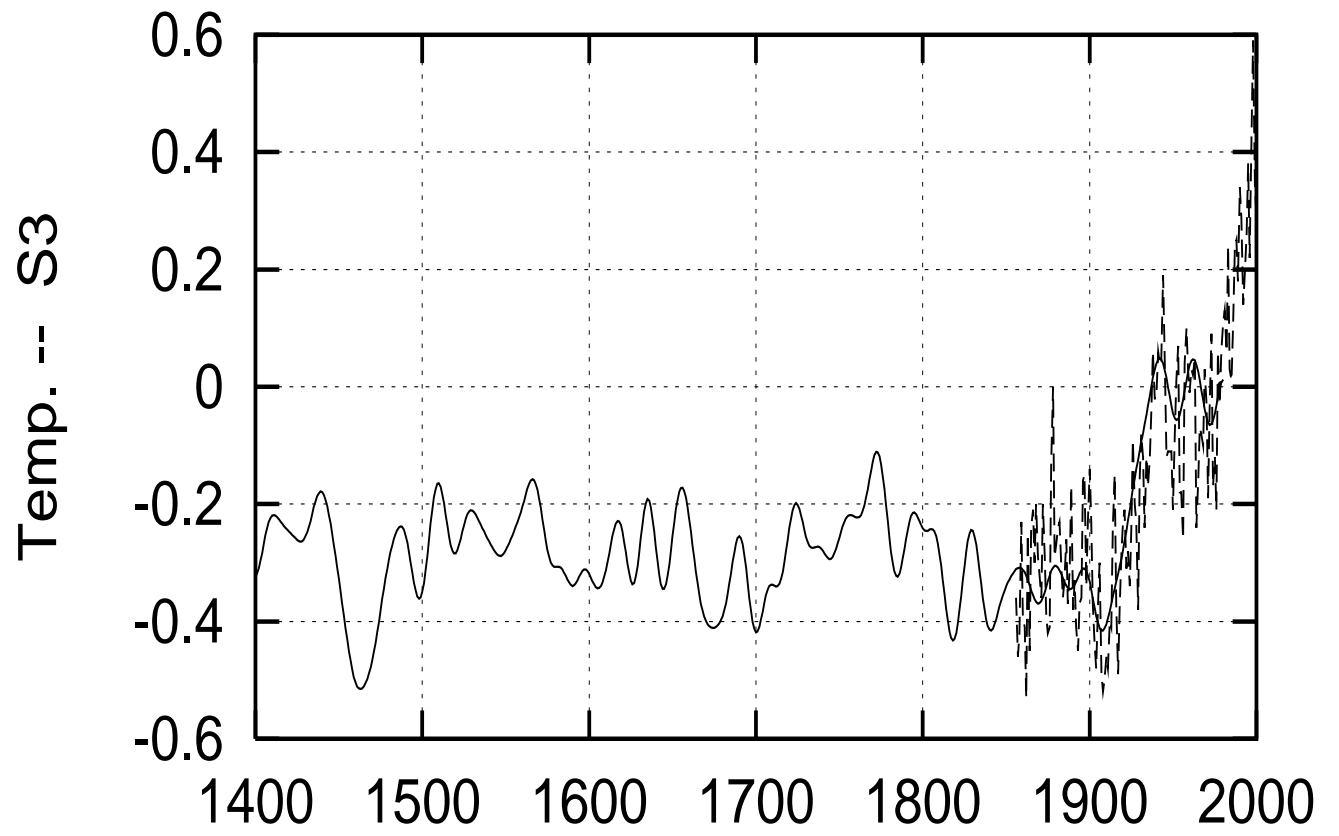


FIG. 5. Mann temperature model: the smooth average S_3 that represent the most accurate smooth reconstruction of the temperature in the year 1400-1980 (solid line). The dashed line is the real temperature in the year 1856-2000.

	r temp vs Briffa	r temp vs Jones	r temp vs Mann
data	0.23	0.56	0.86
S5	0.35	0.88	0.99
S4	0.34	0.84	0.99
S3	0.32	0.80	0.97
S2	0.27	0.72	0.93
S1	0.27	0.64	0.90
D5	0.82	0.76	0.96
D4	0.30	0.17	0.66
D3	0.17	-0.05	0.45
D2	0.30	0.10	0.62
D1	0.20	0.21	0.60

TABLE I. Multiresolution Linear Correlation Analysis